PCT/GB 2004/001253

GB04/001253

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

REC'D 04 MAY 2004
WIPO PCT

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed M Cunkie

Dated 27 April 2004

BEST AVAILABLE COPY

[redacted]

Patents Act 1977
(Rule 16)

THE PATENT OFFICE G
26 MAR 2003
LONDON

The Patent Office

The Patent Office
Cardiff Road
Newport
Gwent NP10 8QQ

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

1.	Your reference	A30289
		26 MAR 2003
2.	Patent application number *(The Patent Office will fill in this part)*	0306973.9
		~~27MAR03 E795436-1 D03052~~ P01/7700 0.00-0306973.9
3.	Full name, address and postcode of the or of each applicant *(underline all surnames)*	BRITISH TELECOMMUNICATIONS public limited company 81 NEWGATE STREET LONDON, EC1A 7AJ, England Registered in England: 1800000
	Patents ADP number *(if you know it)*	~~1867002~~ 6300388001
	If the applicant is a corporate body, give the country/state of its incorporation	UNITED KINGDOM
4.	Title of the invention	TRANSMITTING VIDEO
5.	Name of your agent *(if you have one)*	
	"Address for Service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*	BT Group Legal Intellectual Property Department BT Centre 81 Newgate Street LONDON EC1A 7AJ
	Patents ADP number *(if you know it)*	~~1867001~~ 04..19001
6.	If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number	Country / Priority application number *(if you know it)* / Date of filing *(day / month / year)*
7.	If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application	Number of earlier application / Date of filing *(day/month/year)*
8.	Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:* *a) any applicant named in part 3 is not an inventor, or* *b) there is an inventor who is not named as an applicant, or* *c) any named applicant is a corporate body.* *(See note (d))*	YES

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	-
Description	7
Claim(s)	2
Abstract	-
Drawing(s)	2 +2

10. If you are also filing any of the following, state how may against each item

Priority Documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	1
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. I/We request the grant of a patent on the basis of this application.

Signature(s)

Date: 26 March 2003

LLOYD, Barry George William, Authorised Signatory

12. Name and daytime telephone number of person to contact in the United Kingdom — Rod HILLEN — 020 7492 8140

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

DUPLICATE

Transmitting Video

The present invention is concerned with methods and apparatus for transmitting encoded video material over a network.

According to one aspect of the present invention there is provided a method of transmitting encoded video over a network to a terminal, comprising: storing a plurality of encoded versions of the same video sequence, wherein each version comprises a plurality of discrete portions of data and each version corresponds to a respective different degree of compression; ascertaining the data rate permitted by the network; ascertaining the state of a receiving buffer at the terminal; for each version, computing for discrete portions thereof as yet unsent the value of a timing error that would occur were any number of portions starting with that portion to be sent at the currently ascertained permitted rate; for each version, determining for each of at least some of the discrete portions thereof as yet unsent the maximum of the error values for that portion; for each version, comparing the determined maximum error value with the ascertained buffer state; selecting one of said versions for transmission, in dependence on the results of said comparisons; and
transmitting the selected version.

In another aspect, the invention provides a method of transmitting encoded video over a network to a terminal, comprising: storing a plurality of encoded versions of the same video sequence, wherein each version comprises a plurality of discrete portions of data and each version corresponds to a respective different degree of compression; for each version and for each of a plurality of nominal transmitting rates, computing for discrete portions thereof the value of a timing error that would occur were any number of portions starting with that portion to be sent at the respective nominal rate; for each version and for each of said plurality of nominal transmitting rates, determining for each of at least some of the discrete portions thereof the maximum of the error values for that; storing said maximum error values; ascertaining the data rate permitted by the network; ascertaining the state of a receiving buffer at the terminal; for each version, using the ascertained permitted data rate and the stored maximum error values to estimate a maximum error value corresponding to said ascertained permitted data rate; for each version, comparing the estimated maximum error value with the ascertained buffer state; selecting one of said versions for transmission, in dependence on the results of said comparisons; and transmitting the selected version.

Further aspects of the invention are set out in the claims

Some embodiments of the invention will now be described, by way of example, with reference to the accompanying drawings.

In Figure 1, a streamer 1 contains (or has access to) a store 11 in which are stored files each being a compressed version of a video sequence, encoded using a conventional compression algorithm such as that defined in the ITU standard H.261 or H.263, or one of the ISO MPEG standards. More particularly, the store 11 contains, for the same original video material, several files each encoded with a different degree of compression. In practice all the material could if desired be stored in one single file, but for the purposes of description they will be assumed to be separate files. Thus Figure 1 shows three such files: V1, encoded with a high degree of compression and hence low bit-rate, representing a low-quality recording; V2, encoded with a lesser degree of compression and hence higher bit-rate, representing a medium-quality recording; and V3, encoded with a low degree of compression and hence even higher bit-rate, representing a high-quality recording. Naturally one may store similar multiple recordings of further video sequences, but this is not important to the principles of operation.

By "bit-rate" here is meant the bit-rate generated by the original encoder and consumed by the ultimate decoder; in general this is not the same as the rate at which the streamer actually transmits, which will be referred to as the transmitting bit-rate. It should also be noted that these files are generated at a variable bit-rate (VBR) - that is, the number of bits generated for any particular frame of the video depends on the picture content. Consequently, references above to low (etc.) bit-rate refer to the average bit-rate.

The server has a transmitter 12 which serves to output data via a network 2 to a terminal 3. The transmitter is conventional, perhaps operating with a well known protocol such as TCP/IP. A control unit 13 serves in conventional manner to receive requests from the terminal for delivery of a particular sequence, and to read packets of data from the store 1 for sending to the transmitter 12 as and when the transmitter is able to receive them. Here it is assumed that the data are read out as discrete packets, often one packet per frame of video, though the possibility of generating more than one packet for a single frame is not excluded. (Whilst is in principle possible for a single packet to contain data for more than one frame, this is not usually of much interest in practice).

Note that these packets are not necessarily related to any packet structure used on the network 2.

The terminal 3 has a receiver 31, a buffer 32, primarily for accommodating short-term fluctuations in network delay and throughput, and a decoder 33. In principle, the terminal is conventional, though to get full benefit from the use of the server, one might choose to use a terminal having a larger buffer 32 than is usual.

Some networks (including TCP/IP networks) have the characteristic that the available transmitting data rate fluctuates according to the degree of loading on the network. The reason for providing alternative versions V1, V2, V3 of one and the same video sequence is that one may choose a version that the

network is currently able to support. Another function of the control unit 13, therefore, is to interrogate the transmitter 12 to ascertain the transmitting data rate that is currently available, and take a decision as to which version to send. Here, as in many such systems, this is a dynamic process: during the course of a transmission the available rate is continually monitored so that as conditions improve (or deteriorate) the server may switch to a higher (or lower) quality version. Sometimes (as in TCP/IP) the available transmitting rate is not known until after transmission has begun; one solution is always to begin by sending the lowest-rate version and switch up if and when it becomes apparent that a higher quality version can be accommodated.

Some systems employ additional versions of the video sequence representing transitional data which can be transmitted between the cessation of one version and the commencement of a different one, so as to bridge any incompatibility between the two versions. If required, this may be implemented, for example , in the manner described in our U.S. patent 6,002,440.

In this description we will concentrate on the actual decision on if and when to switch. Conventional systems compare the available transmitting bit-rate with the *average* bit-rates of the versions available for transmission. We have recognised, however, that this is unsatisfactory for VBR systems because it leaves open the possibility that at some time in the future the available transmitting bit-rate will be insufficient to accommodate short-term fluctuations in instantaneous bit-rate as the latter varies with picture content. Some theoretical discussion is in order at this point.

As shown in Figure 2, an encoded video sequence consists of N packets. Each packet has a header containing a time index t_i (i=0 ... N-1) (in terms of real display time – e.g. this could be the video frame number) and contains b_i bits. This analysis assumes that packet *i* must be completely received before it can be decoded arrives (i.e. one must buffer the whole packet first).

In a simple case, each packet corresponds to one frame, and the time-stamps t_i increase monotonically, that is, $t_{i+1} > t_i$ for all i. If however (i) a frame can give rise to two or more packets (each with the same t_i) then $t_{i+1} \geq t_i$. If (ii) frames can run out of capture-and-display sequence (as in MPEG) then the t_i do not increase monotonically

These times are relative. Suppose the receiver has received packet 0 and starts decoding packet 0 at time $t_{ref}+t_0$. At "time now" of $t_{ref}+t_g$ the receiver has received packet t_g (and possibly more packets too) and has just started to decode packet g.

Packets g to h-1 are in the buffer. Note that(in the simple case) if h = g + 1 then the buffer contains packet g only. At time $t_{ref}+t_j$ the decoder is required to start decoding packet j. Therefore, at that time $t_{ref}+t_j$ the decoder will need to have received all packets up to and including packet j.

The time available from now up to $t_{ref} + t_j$ is $(t_{ref} + t_j) - (t_{ref} + t_g) = t_j - t_g$. (1)

The data to be sent in that time are that for packets h to j, viz.

$$\sum_{i=h}^{j} b_i \tag{2}$$

which at a transmitting rate R will require a transmission duration

$$\frac{\sum_{i=h}^{j} b_i}{R} \tag{3}$$

This is possible only if this transmission duration is less than or equal to the time available, i.e. when the currently available transmitting rate R satisfies the inequality

$$\frac{\sum_{i=h}^{j} b_i}{R} \le t_j - t_g \tag{4}$$

Note that this is the condition for satisfactory reception and decoding of frame j: satisfactory transmission of the whole of the remaining sequence requires that this condition be satisfied for all j = h ... N-1.

For reasons that will become apparent, we rewrite Equation (4) as:

$$\frac{\sum_{i=h}^{j} b_i}{R} - (t_j - t_{h-1}) \le t_{h-1} - t_g \tag{5}$$

Note that $t_j - t_{h-1} = \sum_{i=h}^{j} (t_i - t_{i-1}) = \sum_{i=h}^{j} \Delta t_i$ where $\Delta t_i = t_i - t_{i-1}$.

Also, we define $\Delta \varepsilon_i = (b_i / R) - \Delta t_i$

and $T_B = t_{h-1} - t_g$; note that T_B is the difference between the time-stamp of the most recently received packet in the buffer and the time stamp of the least recently received packet in the buffer – i.e. the one that we have just started to decode.

Then the condition is

$$\sum_{i=h}^{j} \Delta \varepsilon_i \le T_B \tag{6}$$

For a successful transmission up to the last packet N-1, this condition must be satisfied for any possible j, viz.

$$Max_{j=h}^{j=N-1}\left\{\sum_{i=h}^{j}\Delta\varepsilon_i\right\} \le T_B \qquad (7)$$

The left-hand side of Equation (7) represents the maximum timing error that may occur from the transmission of packet h up to the end of the sequence, and the condition states, in effect that this error must not exceed the ability of the receiver buffer to accommodate it, given its current contents. For convenience, we will label the left-hand side of Equation (7) as T_h - i.e.

$$T_h = Max_{j=h}^{j=N-1}\left\{\sum_{i=h}^{j}\Delta\varepsilon_i\right\} \qquad (8)$$

In practice we prefer to allow switching only at certain defined "switching points" in the sequence (and naturally provide the transitional data mentioned earlier only for such points). In that case the test needs to be performed only at such points. The switching decision at frame h may proceed as follows:

- interrogate the transmitter 12 to determine the available transmitting rate R;
- ascertain the current value of T_B: this may be calculated at the terminal and transmitted to the server, or may be calculated at the server (see below;
- compute (for each file V1, V2, V3) T_h in accordance with Equation (8) - let these be called $T_h(1)$, $T_h(2)$, $T_h(3)$;
- determine the highest value of k for which $T_h(k) + \Delta \le T_B$, where Δ is a fixed safety margin;
- select file V_k for transmission.

The calculation of T_B at the server will depend on the exact method of streaming that is in use.

Our preferred method is (as described our in international patent application no. PCT/GB 01/05246 [Agent's Ref. A26079]) to send, initially, video at the lowest quality, so that the terminal may immediately start decoding whilst at the same time the receiving buffer can be filling up because data is being sent at a higher rate than it is used. In this case the server can deduce current client session time (i.e. the timestamp of the packet currently being decoded at the terminal) without any feedback, and so

T_B = latest sent packet time – current client session time.

If the system is arranged such that the terminal waits until some desired state of buffer fullness is reached before playing begins, then the situation is not quite so simple because there is an additional delay to take into account. If this delay is fixed, it can be included in the calculation. Similarly, if the

terminal calculates when to start playing and both the algorithm used, and the parameters used by the algorithm, are known by the server, again this can be taken into account. If however the terminal is of unknown type, or controls its buffer on the basis of local conditions, feedback from the terminal will be needed.

Now, this procedure will work perfectly well, but does involve a considerable amount of processing that has to be carried out during the transmission process. In a modified implementation, therefore, we prefer to perform as much as possible of this computation in advance. In principle this involves the calculation of $T_h(k)$ for every packet that follows a switching point, and storing this value in the packet header. Unfortunately, this calculation (Equation (8) and the definition of $\Delta\varepsilon_i$) involves the value of R, which is of course unknown at the time of this pre-processing. Therefore we proceed by calculating $T_h(k)$ for a selection of possible values of R, for example (if R_A is the average bit rate of the file in question)

$$R_1 = 0.5R_A$$
$$R_2 = 0.7R_A$$
$$R_3 = R_A$$
$$R_4 = 1.3R_A$$
$$R_5 = 2R_A$$

So each packet h has these five precalculated values of T_h stored in it. If required (for the purposes to be discussed below) one may also store the relative time position at which the maximum in Equation (8)) occurs, that is,

$\Delta t_{h\max} = t_{j\max} - t_h$ where $t_{j\max}$ is the value of j in Equation 8 for which T_h is obtained.

In this case the switching decision at frame h proceeds as follows:

- interrogate the transmitter 12 to determine the available transmitting rate R;
- ascertain the current value of T_B, as before;
- EITHER - in the event that R corresponds to one of the rates for which T_h has been precalculated - read this value from the store (for each file V1, V2, V3);
- OR - in the event that R does not so correspond, read from the store the value of T_h (and, if required, $t_{h\max}$) that correspond to the highest one (R^-) of the rates $R_1 \ldots R_5$ that is less than the actual value of R, and estimate T_h from it (again, for each file V1, V2, V3);
- determine the highest value of k for which $T_h(k) + \Delta \leq T_B$, where Δ is a fixed safety margin;
- select file V_k for transmission.

The estimate of T_h could be performed simply by using the value T_h^- associated with R^-; this would work, but since it would overestimate T_h it would result, at times, in a switch to a higher quality stream being judged impossible even though it were possible. Another option would be by linear (or other) interpolation between the values of T_h stored for the two values of R_1 ... R_5 each side of the actual value R. However, our preferred approach is to calculate an estimate according to:

$$T_i' = \frac{(T_i^- + \Delta T^-_{i\max})R^-}{R} - \Delta T^-_{i\max}$$

Where R^- is the highest one of the rates R_1...R_5 that is less than the actual value of R, T_i^- is the precalculated T_h for this rate, $\Delta t^-_{i\max}$ is the time from t_i at which T_i^- is obtained (i.e. is the accompanying value of $\Delta t^-_{h\max}$ .. In the event that this method returns a negative value, we set it to zero.

Note that this is only an estimate, as T_h is a nonlinear function of rate. However with this method t_i' is always higher than the true value and automatically provides a safety margin (so that the margin Δ shown above may be omitted.

Note that these equations are valid for the situation where the encoding process generates two or more packets (with equal t_i) for one frame, and for the situation encountered in MPEG with bidirectional prediction where the frames are transmitted in the order in which they need to be decoded, rather than in order of ascending T_i.

Claims

1. A method of transmitting encoded video over a network to a terminal, comprising storing a plurality of encoded versions of the same video sequence, wherein each version comprises a plurality of discrete portions of data and each version corresponds to a respective different degree of compression;
ascertaining the data rate permitted by the network;
ascertaining the state of a receiving buffer at the terminal;
for each version, computing for discrete portions thereof as yet unsent the value of a timing error that would occur were any number of portions starting with that portion to be sent at the currently ascertained permitted rate;
for each version, determining for each of at least some of the discrete portions thereof as yet unsent the maximum of the error values for that portion;
for each version, comparing the determined maximum error value with the ascertained buffer state;
selecting one of said versions for transmission, in dependence on the results of said comparisons; and transmitting the selected version.

2. A method of transmitting encoded video over a network to a terminal, comprising storing a plurality of encoded versions of the same video sequence, wherein each version comprises a plurality of discrete portions of data and each version corresponds to a respective different degree of compression;
for each version and for each of a plurality of nominal transmitting rates, computing for discrete portions thereof the value of a timing error that would occur were any number of portions starting with that portion to be sent at the respective nominal rate;
for each version and for each of said plurality of nominal transmitting rates, determining for each of at least some of the discrete portions thereof the maximum of the error values for that;
storing said maximum error values;
ascertaining the data rate permitted by the network;
ascertaining the state of a receiving buffer at the terminal;
for each version, using the ascertained permitted data rate and the stored maximum error values to estimate a maximum error value corresponding to said ascertained permitted data rate;
for each version, comparing the estimated maximum error value with the ascertained buffer state;

selecting one of said versions for transmission, in dependence on the results of said comparisons; and transmitting the selected version.

3. A method according to claim 1 or 2 in which said maximum timing error determination is performed only for selected ones of said portions at which a version change is to be permitted.

4. A method according to claim 1, 2 or 3 in which each computed timing error value is the difference between (a) the time needed to transmit, at the relevant rate, the portion in question and zero or more consecutive subsequent portions up to and including any particular portion, and (b) the difference between the playing instant of the respective particular portion and the playing instant of the portion preceding the portion in question.

5. A video recording stored on a carrier, comprising

a plurality of encoded versions of the same video sequence, wherein each version comprises a plurality of discrete portions of data and each version corresponds to a respective different degree of compression; and

for each discrete portion of each version and for each of a plurality of nominal transmitting rates, a maximum error value for that portion, being the maximum of (a) the value of a timing error that would occur were that portion to be sent at the respective nominal rate; and

(b) the values of a timing error that would occur were that portion and any number of subsequent portions subsequent thereto to be sent at the respective nominal rate.

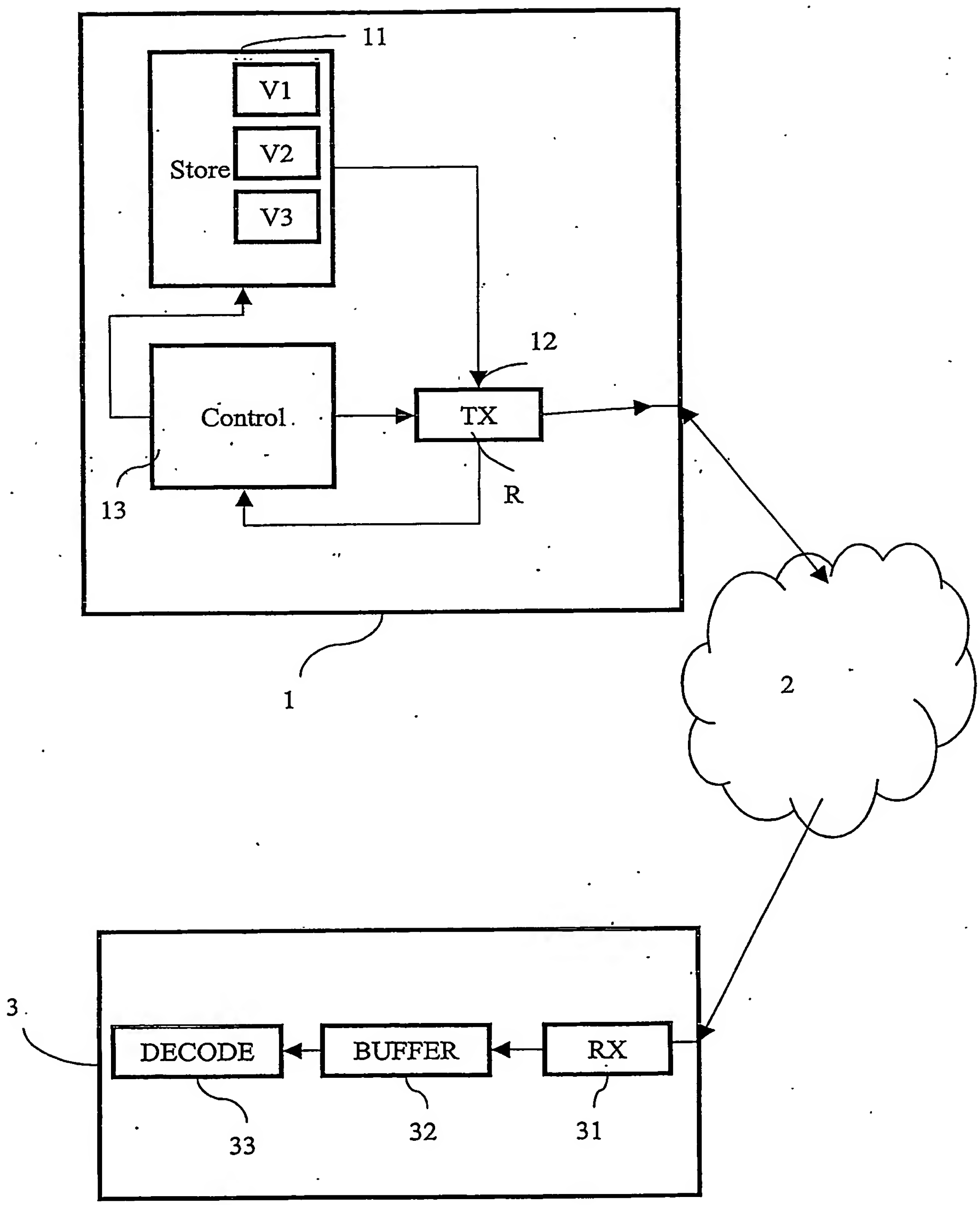
11
V1
Store
V2
V3
12
Control
TX
R
13
1
2
3
DECODE
BUFFER
RX
33
32
31

Figure 1

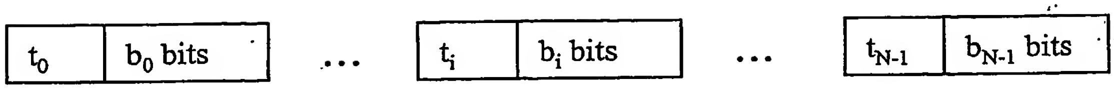

Figure 2